

AD-A080 851

ARIZONA STATE UNIV TEMPE

F/G 5/10

ATTENTION AND TASK COMPLEXITY AS INDICATED BY PHYSIOLOGICAL IND--ETC(U)

DEC 79 E LINDHOLM, M RUPPEL, S H BUCKLAND

F33615-77-C-0069

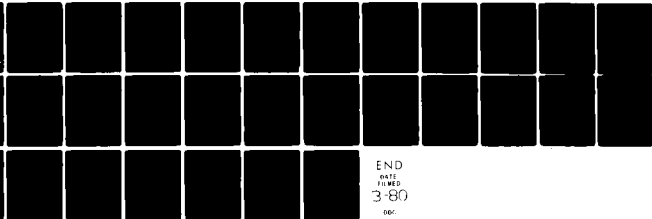
NL

UNCLASSIFIED

AFHRL-TR-79-47

1 OF 1

AD
ADDRESS:



END
DATE
10-80
3-80

DEC:

AFHRL-TR-79-47

AIR FORCE



ADA 080851

DDC FILE COPY

HUMAN RESOURCES

LEVEL II

ATTENTION AND TASK COMPLEXITY AS INDICATED
BY PHYSIOLOGICAL INDICES

By

Ernest Lindholm
Michael Ruppel
Arizona State University
Tempe, Arizona 85281

2

George H. Buckland, Capt, USAF

FLYING TRAINING DIVISION
Williams Air Force Base, Arizona 85224

December 1979

Final Report

DDC
RECEIVED
FEB 14 1980
E

Approved for public release; distribution unlimited.

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 76125

80 2 11 007

NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Arizona State University, Tempe, Arizona 85281, under contract F33615-77-C-0069, project 2313, with Flying Training Division, Air Force Human Resources Laboratory (AFSC), Williams Air Force Base, Arizona 85224. Capt George H. Buckland (FTR) was the Principal Investigator for the Laboratory.

This report has been reviewed by the Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

DIRK C. PRATHER, Lieutenant Colonel, USAF
Technical Advisor, Flying Training Division

RONALD W. TERRY, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| 17 REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|--|---|
| 1. REPORT NUMBER AFHRLTR-79-47 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) ATTENTION AND TASK COMPLEXITY AS INDICATED BY PHYSIOLOGICAL INDICES | 5. TYPE OF REPORT & PERIOD COVERED Final rept. | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) Ernest Lindholm Michael Ruppel George H. Buckland | 8. CONTRACT OR GRANT NUMBER(s) F33615-77-C-0069 | |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Arizona State University Tempe, Arizona 85281 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 23137509 | |
| 11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235 | 12. REPORT DATE Dec 1979 | |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Flying Training Division Air Force Human Resources Laboratory Williams Air Force Base, Arizona 85224 | 13. NUMBER OF PAGES 34 | |
| | 15. SECURITY CLASS. (of this report) Unclassified | |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE | |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) attention psychophysiology evoked potentials reaction time galvanic skin response task complexity heart rate workload human information processing | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The objective of this research was to develop physiological measures of pilot attention and workload for use in flight simulation research studies. In later studies, these physiological measures will then be used in conjunction with behavioral measures of pilot attention and task difficulty in order to optimally structure flight simulation training programs and equipment. This research employed two common behavioral information processing tasks in order to study four physiological variables, namely, heart rate, skin conductance, eye movement, and cortical evoked response, as measures of attention and arousal. Four separate information processing experiments were conducted involving both choice reaction time and letter matching tasks. For the choice reaction time tasks, the faster responses were in general associated with increased measures of arousal as indicated by heart rate, skin conductance, | | |

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Item 20 Continued:

and evoked potential amplitude. There was no general arousal effect produced for the letter matching tasks; however, the cortical evoked response varied systematically and reliably with task difficulty and reaction time. In this regard, simple decisions evoked short latency low amplitude brain waves, while more difficult decisions evoked longer latency, higher amplitude brain waves. Thus, the heart rate and skin conductance measures of arousal reflected, in a gross sense, the degree of subject involvement in the task, and the evoked potential components varied reliably with task difficulty and performance.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Preface

The authors thank Lt. Col. Robert O'Donnell, Wright-Patterson Air Force Base, for sharing his broad and expert knowledge of all phases of cortical evoked potential research. Gratitude is also extended to Dr. Stanley R. Parkinson, Arizona State University, for methodological advice and to Dr. Julianne M. Lindholm, Arizona State University, who served as statistical advisor. Special thanks are extended to Mr. Gary Reid, Williams Air Force Base, who followed all phases of this project with sincere interest and critical appraisal.

| | |
|--------------------|--|
| Accession For | |
| NTIS GRA&I | <input checked="checked" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | <input type="checkbox"/> |
| By | |
| Distribution/ | |
| Availability Codes | |
| Dist | Avail and/or special |
| A | |

TABLE OF CONTENTS

| | Page |
|------------------------------------|------|
| I. Introduction | 1 |
| II. Rationale. | 2 |
| III. Objectives | 2 |
| IV. Experiment 1 | 3 |
| Introduction. | 3 |
| Method. | 3 |
| Results and Discussion. | 5 |
| V. Experiment 2 | 12 |
| Introduction. | 12 |
| Method. | 12 |
| Results and Discussion. | 12 |
| VI. Experiment 3 | 16 |
| Introduction. | 16 |
| Method. | 17 |
| Results and Discussion. | 18 |
| VII. Experiment 4 | 21 |
| Introduction. | 21 |
| Method. | 21 |
| Results and Discussion. | 22 |
| VIII. General Discussion | 24 |
| References. | 26 |

LIST OF ILLUSTRATIONS

| Figure | Page |
|--|------|
| 1 Heart rate shown in top two panels and skin conductance shown in bottom two panels as functions of bits of information and trial blocks for the fast and slow subjects of Experiment 1 | 7 |
| 2 Vertex potentials as functions of bits of information for the fast and slow subjects of Experiment 1 | 8 |

List of Illustrations (Continued)

| Figure | | Page |
|--------|---|------|
| 3 | Left vertical mark is both calibration marker and stimulus onset marker. Positive up. Each tic on abscissa is 100 msec. Broken line is vertex potential for fast (Q1) RTs, and solid line is vertex potential for slow(Q4) RTs. | 10 |
| 4 | Heart rate shown in top two panels and skin conductance shown in bottom two panels as functions of bits of information and trial blocks for the fast and slow subjects of Experiment 2. | 13 |
| 5 | Vertex potentials as functions of bits of information for the fast and slow subjects of Experiment 2 . . . | 14 |
| 6 | Left vertical mark is both calibration marker and stimulus onset marker. Positive up. Abscissa tics are 100 msec apart. Light traces are vertex potentials for fast RTs, heavy traces, slow RTs. Top traces, Condition 1, middle traces, Condition 2, and bottom traces, Condition 3.. . . . | 15 |
| 7 | Left vertical mark is both calibration and stimulus onset marker. Positive up. Abscissa tics are 100 msec apart. Heavy solid trace is vertex potential for PI matches, light solid trace is vertex potential for NI matches, and broken trace is vertex potential for VC matches | 20 |
| 8 | Left vertical mark is both calibration and stimulus onset marker. Positive up. Abscissa tics are 100 msec apart. Top traces, vertex potential responses to first letter of pair. Bottom traces, vertex potential responses to second letter of pair presented 900 msec after first letter. Light traces are NI matches; heavy traces, PI matches. | 23 |

ATTENTION AND TASK COMPLEXITY AS INDICATED BY PHYSIOLOGICAL INDICES

I. Introduction

Traditional assessment of human performance emphasizes the measurement of one or another skeletal motor responses, for example, writing or speaking answers to questions or moving the limbs and digits to manipulate knobs, levers, or switches. Dependent variables are quantified in terms of percentage correct responses, speed of response (reaction time (RT)), RMS error, time-on-target, and other similar metrics. Recently developed techniques, however, enable the assessment of the internal state of the operator performing the task. This opens a new and rich area of research in which human performance can be quantitatively related to variables such as arousal and attention.

The three psychophysiological variables receiving the most attention by investigators are heart rate, skin conductance, and the cortical evoked potential (EP). This trio of responses is particularly informative since heart rate (HR) and the skin conductance response (SCR) vary systematically with the level of sympathetic arousal and both are components of the orienting response. The cortical evoked response reflects central nervous system events rather than autonomic nervous system functions and can be used to investigate attentional states and mental effort. All three variables have been empirically related to task difficulty. HR varies systematically with effort in a variety of tasks including mental arithmetic, ski jumping (Gunn, Wolf, Block, & Person, 1972), making psychophysical judgments (Lang, Gatchel, & Simons, 1975), RT tasks (Lacey & Lacey, 1978), and combat flying (Roman, Older & Jones, 1967). SCR changes more slowly than HR and is useful for determining tonic changes in sympathetic activity which occur over periods of a half hour or so. Several years of research have related SCR to physiological and psychological activation, emotion, learning, stress, personality traits, and pathological states (Edelberg, 1972).

The cortical evoked potential (EP, also called the "event related potential" (ERP)), is a complex phenomenon. Following the presentation of a discrete visual or auditory stimulus, the electrical activity of the brain undergoes a series of positive and negative oscillations. Each plus and minus swing of the brain wave is referred to as a "component" of the EP and the components of primary interest to behavioral researchers take place within the first second following stimulus onset. Two classic experiments related changes in EPs to vigilance and selective attention. Haider, Spong, and Lindsley (1964) showed that certain components of the EP varied reliably with performance on a vigilance task, suggesting that examination of the EP waveform could be used to infer level of attention. Regarding selective attention, Spong, Haider, and Lindsley (1965)

showed that, if subjects were instructed to attend to one sensory channel and ignore another, EPs to the attended channel stimuli were augmented.

In the tradition of the vigilance paradigm much of the recent EP research has focused on responses to rare events, e.g., an infrequent high tone occurring in a series of low tones (Karlin, 1970; Sutton, 1969; Goodin, Squires, Henderson, & Starr, 1978; Duncan-Johnson & Donchin, 1977).

While the three variables of HR, SCR, and EP are conceptually related to each other and to task difficulty, there has been a surprising lack of research aimed at interrelating the three. This report discusses four experiments designed to investigate the extent to which HR, SCR, and EP co-vary with each other and with performance during two common behavioral information processing tasks. The tasks are choice reaction time and letter matching, chosen because of their fundamental importance to the study of human information processing. Two versions of each task are conducted, and within each task, information load is systematically varied.

II. Rationale

Assessment of human performance becomes more difficult as the complexity of man-machine systems increases. The point has been reached where behavioral research must step beyond the limits imposed by quantifying behavior in terms of skeletal motor responses alone. Psychophysiological assessment of the internal state of the operator shows promise of providing the tools to take this step. By combining behavioral and psychophysiological assessment, a more comprehensive profile of human performance emerges. This should permit a greater understanding of the conditions under which performance deteriorates and point to techniques which can maximize performance.

III. Objectives

The project discussed in this report utilizes a newly established computer laboratory. One primary objective was to establish laboratory procedures for efficient and reliable handling of the large quantities of data involved in psychophysiological research and to insure that all components of the computer laboratory were functioning properly. This objective included the training of project personnel in the use of the new computer facility.

The variables of HR, SCR, EP, and eye movement have not been investigated simultaneously in previous research. A major objective was to investigate the interrelationships of these variables with each other and with performance in two common

information processing tasks.

IV. Experiment 1

Introduction

Although RT tasks have been used previously in psychophysiological research, the paradigms have differed considerably from the classic choice RT paradigm. Particularly in EP research, investigators tend to use simple RT rather than choice RT, and the subjects are instructed to respond to an infrequent signal, usually a tone. Operationally, then, these paradigms become vigilance tasks. Also, motor movements are minimized by making the RT response as simple as possible, thus there is little generality to real-life situations where responses involve arm, wrist, and finger movements. Finally, as mentioned before, investigators typically do not quantify several psychophysiological variables simultaneously and attempt to interrelate them.

The purpose here was to simulate a more real-life situation, as, for example, in a cockpit, where one of several lights might come on, each one requiring a different response. For this reason, the classic 8-alternative choice RT paradigm was used, in which RT typically increases linearly as a function of bits of information (log of the number of alternatives).

Method

Subjects-Six male psychology graduate students (ages 21 to 29 years, mean of 25 years) were paid \$4.00 per hour for participation. All were familiar with choice RT paradigms but none had previously served as a subject in such a paradigm.

Apparatus-The RT panel consisted of eight response buttons arranged in a semicircle with a start button in the center. Each button was 2.5 cm square and could be independently lighted. The radius of the semicircle, measured from the start button to any response button, was 20.5 cm. Since the response buttons could be independently lighted, they also served as the stimulus; this arrangement maximized stimulus-response compatibility.

The RT panel rested on a shelf placed 68 cm above the floor of an electrically shielded subject booth, which had inside dimensions of 1.2 meters long by 0.8 meter wide by 1.7 meters high. The booth contained only the RT panel and a common armless steel-frame, vinyl padded chair in which the subject sat.

Biopotentials were led from Beckman silver-silver chloride electrodes to a Beckman Type 411 6-channel Dynograph, the high level outputs of which were fed to a Digital Equipment Corporation PDP 11/34 computer fitted with Analog-to-digital (A/D) converters. The PDP 11/34 system was programmed to control all phases of the experiment. Data collected on-line were stored on a Perdec magnetic disk and were then transferred to magnetic tape for storage and off-line analysis. Reaction time was measured to the nearest millisecond by a programmable hardware clock. The A/D was programmed to sample and store each 4 msec.

Procedure-A trial began when computer software lighted the rest button. The subject depressed the rest button which extinguished the rest button light and initiated a variable foreperiod (3 to 6 sec, mean of 4.5 sec) at the end of which one response button was lighted. In the 2-alternative condition, one of the two topmost buttons was lighted. In the 4-alternative condition, one of the four topmost buttons was lighted, and in the 8-alternative condition, any one of the eight response buttons was lighted. Which button was lighted was determined randomly with the constraints that runs greater than three were not permitted and that stimulus probability was equated over a 100 trial sequence.

The subject's task was simply to release the start button and press the lighted button as rapidly as possible. This extinguished the light and initiated a fixed 4 sec intertrial interval after which the start button was again lighted to signal that the next trial could begin. Subjects were cautioned to restrict gross body movements to the intertrial interval and to avoid pressing the start button until they were ready to remain motionless. In this manner, the subject could prolong the intertrial interval in order to make postural adjustments and thus avoid movement artifacts in the physiological records during a trial.

If the start button was pressed and then released during the foreperiod, the light came on again, the trial was aborted, and the foreperiod recycled. This insured that all subjects had their finger at the same starting location on every trial. Finally, a "bad" trial was defined as (a) pressing a response button that was not lighted (error) or (b) the failure to press any response button within 1000 msec after one response button was lighted (no response). Bad trials occurred rarely, but when they did, the software compensated by administering another trial so that 100 "good" trials were obtained from each subject in each condition.

There were two sessions separated by 1 to 3 days. During the first session (practice), the task was carefully explained

and a few trials demonstrated. The condition (2, 4, or 8 alternative) was announced and 100 good trials administered at each of the three conditions with 5 min breaks between conditions. The order of presentation of conditions was counterbalanced across subjects. Total time for the practice session was approximately 1 hour.

The second session was identical to the first except that physiological responses were assessed. Electrodes were placed at the following locations: Vertex (C_z) referenced to right mastoid to record the EP; left index finger referenced to the back of the same hand to assess skin conductance; centerline lower rib referenced to lateral line lower rib to assess heart rate; lateral canthus of the left eye referenced to the mediolateral center of the same eye just superior to the eyebrow, to assess eye movements. The left mastoid served as the patient ground. Bandpass was set at 5.3 Hz to 30 Hz for all channels except SCR which was DC to 30 Hz. Accuracy of amplitude measurements was insured by equating Dynograph calibration signals with a known number of A/D units at the beginning of each run.

The subject wore a lightweight junction box around the neck which served to connect the primary leads to the Dynograph through a connector. In this manner the subject could disconnect from the Dynograph during rest periods and move about freely. Electrode impedance (measured at 30 Hz) was typically 1 to 5 kilohms for the two mastoids and vertex leads, and 10 to 40 kilohms for the other leads. All subjects were right-handed and responded with that hand. The left hand rested on a cushion next to the RT panel. Because of the time required to attach and detach the electrodes, this session was approximately 45 min longer than the first.

The experiment was conducted in a darkened room (overhead lights off) so that stimulus onset was a very salient event.

Results and Discussion

In all statistical tests, a .05 rejection level was used to indicate a reliable difference.

Data Editing-Trials were discarded if large ocular responses (blinks) contaminated the vertex potential record. This occurred in less than 5% of the trials.

Behavioral Results-The median RT was computed for each subject and Condition (Conditions are designated 1, 2, and 3 representing bits of information for the 2-, 4-, and 8-alternative choices, respectively). The means of these medians were 344, 340, and 345 msec for Conditions 1, 2,

and 3. This function has a slope of essentially zero, indicating that RT did not increase as a function of bits of information. This result was unexpected; in retrospect, it was hypothesized that the high stimulus-response compatibility and the easily discriminable visual cue combined to make the task extremely easy. Even in the 8-alternative situation, the task required minimal central processing, since it was only necessary for the subject to respond to any abrupt change in background illumination. This hypothesis was tested and supported in Experiment 2.

Psychophysiological Results: Between-Subject Comparisons-As with the RT results, there were no consistent differences among the psychophysiological measures attributable to Conditions. However, interesting relationships emerged when subjects were grouped into fast and slow responders (a simple median split, placing three subjects below and three subjects above the grand median RT). Figure 1 shows the HR and SCR for the fast and slow subjects as functions of Conditions and trials. Considering first the HR results (top two panels), across trials and Conditions, fast subjects maintained higher mean HR than slow subjects. This difference was not reliable ($F(1,4) = 5.73$), although it did approach the .05 level. The groups by Conditions interaction was reliable ($F(2,8) = 5.19$), indicating that the differences in HR between fast and slow subjects decreased as a function Conditions (see Figure 1 top left panel).

The SCR data, displayed in the bottom two panels of Figure 1, were more variable. While inspection of the figure suggests that fast subjects generally maintained higher SCR than slow subjects over Conditions and trials, the only reliable effect involving groups was the groups by trials interaction ($F(2,8) = 7.16$), indicating that fast subjects increased SCR over trials while the slow subjects did not (see bottom right panel of Figure 1).

The vertex potential (i.e., the cortical evoked potential recorded from the vertex lead) was characterized by four prominent peaks; the first negative occurring approximately 200 msec following stimulus onset (N200), a positive at approximately 250 msec (P250), a second negative at approximately 300 msec (N300) and a second positive at approximately 325 msec (P325). As with HR and SCR, a between-subjects analysis was performed on the latency and amplitude of each of these vertex components. Only the N200 and P325 components varied as a function of groups, and these are displayed in Figure 2. N200 amplitude decreased over Conditions for the fast subjects and increased over Conditions for the slow subjects (top panel). This interaction approached, but did not reach, the .05 level of confidence ($F(2,8) = 4.36$).

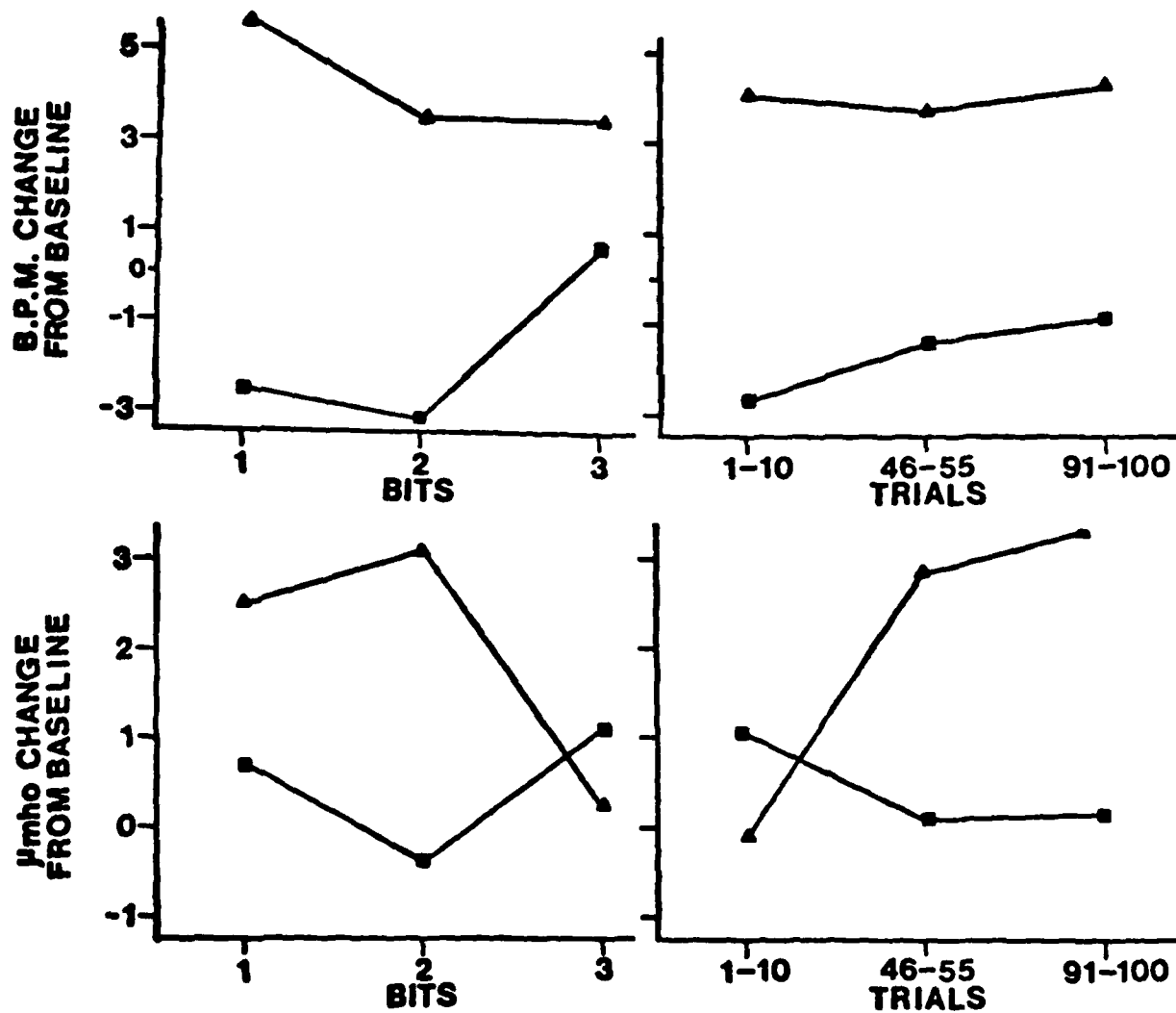


Figure 1. Heart rate (expressed in beats per minute, B.P.M.) shown in top two panels and skin conductance (expressed in μmho) shown in bottom two panels as functions of bits of information (1, 2, and 3 bits correspond to 2, 4, and 8 alternative choice, respectively) and trial blocks for the fast (triangles) and slow (squares) subjects of Experiment 1.

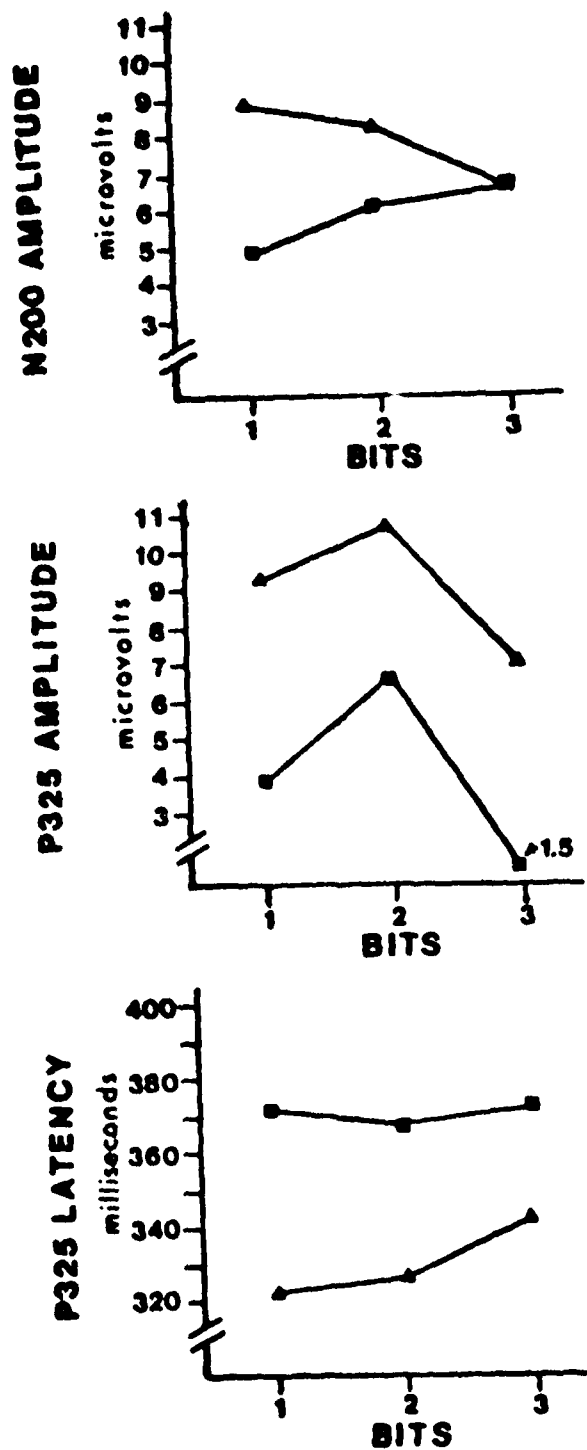


Figure 2. Vertex potentials as functions of bits of information for the fast (triangles) and slow (squares) subjects of Experiment 1.

Similarly, P325 amplitude (middle panel) was apparently greater for fast subjects across Conditions, but the groups effect ($F(1,4) = 6.53$) did not reach the accepted confidence level. Finally, P325 latency (bottom panel) appeared shorter for fast subjects, but not reliably so ($F(1,4) = 3.02$).

In summary, while inspection of the figures suggests that fast subjects were more autonomically aroused (higher HR and SCR) and more attentive (augmented vertex potentials) than slow subjects, statistical support was lacking in several critical comparisons.

Psychophysiological Results: Within-subject comparisons (the between-subject comparisons just discussed necessitated dividing the six subjects into fast and slow subgroups of only three subjects each). While the graphic representation of the results suggested a consistent pattern of augmented physiological responses for fast subjects in all comparisons, statistical support for this generalization was not convincing. Additionally, the problems associated with generalizing from such small sample sizes are well known. The question of whether "good" performance (fast responding) is associated with augmented physiological responses can be approached also by a within-subject analysis which avoids the problem of dividing subjects into subgroups. To pursue this, the RIs for each subject and each Condition were divided into quartiles, then the vertex potential data were averaged separately for Q1 (fastest RIs) and Q4 (slowest RIs). These were analyzed separately for the three Conditions. There were no consistent effects attributable to Conditions, thus for the purposes of graphic representation the three Conditions are averaged and displayed in Figure 3.

The general impression gained from inspection of Figure 3 is that the major components of the vertex potential occur earlier in time and/or are of larger amplitude on fast trials than on slow trials. Statistical analyses lend confidence to this impression: N200 amplitude (from baseline) was greater for fast trials in all three Conditions ($t(5) = 4.69, 2.65, \text{ and } 2.88$ for Conditions 1, 2, and 3, respectively), N300 latency was shorter for fast trials in all three Conditions ($t(5) = 3.00, 4.20, \text{ and } 3.05$), and P325 latency was shorter for fast trials in all three Conditions ($t(5) = 4.79, 6.87, \text{ and } 4.62$). P325 amplitude tended to be larger for fast trials in all three Conditions, but not reliably so ($t(5) = 2.10, 2.55, \text{ and } 1.62$). Thus, the vertex potential was augmented on fast trials, displaying generally shorter latency, higher amplitude components. It appears, therefore, that augmented vertex potentials are associated with fast responding in both within and between subject comparisons.

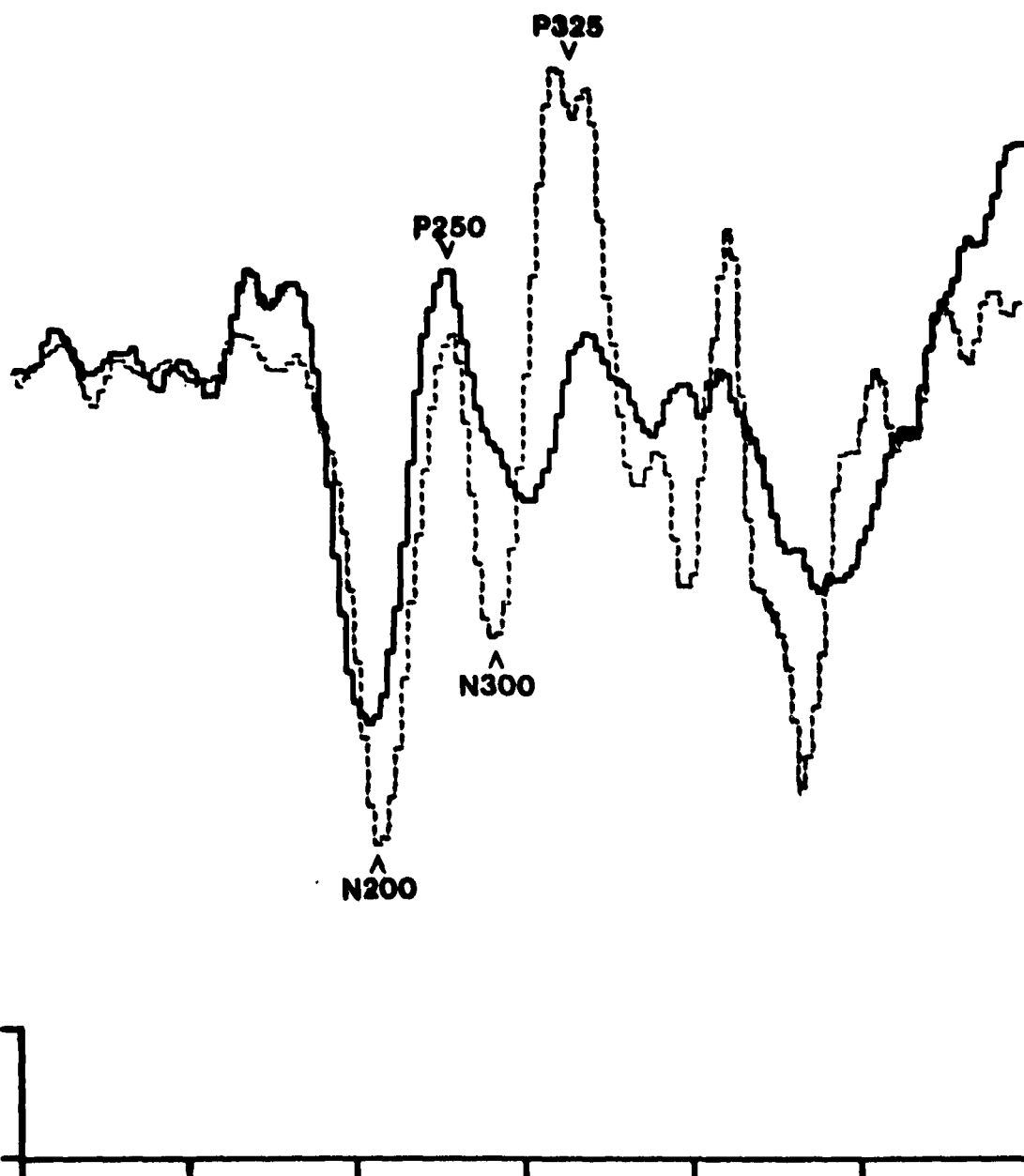


Figure 3. Left vertical mark is both calibration marker (1 uV) and stimulus onset marker. Positive up. Each tic on abscissa is 100 msec. Broken line is vertex potential for fast (Q1) R1s, and solid line is vertex potential for slow (Q4) R1s.

It would be interesting to know whether the HR and SCR also showed reliable differences on fast and slow trials averaged over subjects. Unfortunately, it was necessary to store these data in block format, and it was not possible to sort HR and SCR according to individual trial RI.

Correlational analyses-Within and between subject correlations (Pearson's "r") were computed for RI and latency and amplitude of each of the four prominent components of the vertex potential. The t-test was used to evaluate differences from zero. Considering first the within-subject results; in Conditions 1 and 2, RI correlated reliably with P325 amplitude (average $r = -.28$ and $-.30$ for Conditions 1 and 2) and P325 latency (.25 and .26 for Conditions 1 and 2). The results for Condition 3 were in the same direction but not statistically reliable (.15 for RI and P325 latency, and $-.12$ for RI and P325 amplitude). Thus, in the two simpler conditions, fast RIs were associated with high amplitude, short latency P325s but these relationships deteriorated in the most difficult condition.

The between-subject correlations agreed with the within-subject correlations. RI and P325 amplitude were negatively correlated ($-.47$, $-.13$, and $-.58$ for Conditions 1, 2, and 3), and RI and P325 latency were positively correlated (.48, .59, and .71 for Conditions 1, 2, and 3). With only 4 degrees of freedom available, none of these coefficients were reliably different from zero.

Summary-The data were analyzed from both within and between subject approaches. The between-subject analyses presented a consistent picture of augmented vertex potentials, HR, and SCR suggesting that fast subjects were more sympathetically aroused and more attentive to task-relevant stimuli than were slow subjects, and the correlations suggested that fast RIs were associated with high amplitude, short latency P325s. However, the results of statistical tests in the between analyses often were not statistically reliable. The within central tendency analyses presented the same picture of augmented vertex potentials for fast trials, compared with slow trials, and the correlations again indicated that fast RIs were associated with high amplitude, short latency P325s. In these comparisons, statistical confirmation was reasonably consistent. Thus, the results lead to the conclusion that fast responding is associated with augmented sympathetic and central nervous system activation. This conclusion is examined further in Experiment 2.

V. Experiment 2

Introduction

None of the dependent variables in Experiment 1 varied systematically with Conditions (information load), the major independent variable. Suspecting that this was due to the high stimulus-response (S-R) compatibility in Experiment 1, the present Experiment was designed. In all respects, this Experiment was identical to Experiment 1 with the single exception that central processing demands were increased by decreasing S-R compatibility. Specifically, subjects were instructed to respond, not to the lighted button but to the mirror image of the lighted button. Thus, for example, in the 8-alternative Condition, if the far left button was lighted, the correct response would be to press the far right button even though it was not lighted.

Method

The same six subjects of Experiment 1 served in the present Experiments. All methods and procedures were the same with the single exception of requiring the mirror image responses.

Results and Discussion

Behavioral Results-The means of the subjects' median RTs were 370, 406, and 440 msec for Conditions 1, 2, and 3, respectively. The effect of Conditions was reliable ($p < .05$), indicating that RT increased as a function of bits of information. The function was linear with a slope of 35 msec per bit.

Psychophysiological Results: Between-subject comparisons-The HR and SCR data are displayed in Figure 4. In agreement with the results of Experiment 1, HR and SCR tended to be elevated across Conditions and trials for fast subjects, but none of the main effects or interactions were reliable.

Vertex potential latency and amplitude results for fast and slow subjects are displayed in Figure 5. Neither the main effects nor the interactions were reliable for the N200 amplitude and P325 latency comparisons, but P325 amplitude (middle panel) was reliably greater for fast subjects ($F(1,4) = 7.78$).

Psychophysiological Results: Within-subject comparisons-The vertex potentials for Q1 and Q4 RTs for the three Conditions are displayed in Figure 6. Note that, as task difficulty increased, the P325 component became less well defined, particularly in Condition 3. Examination of the data revealed

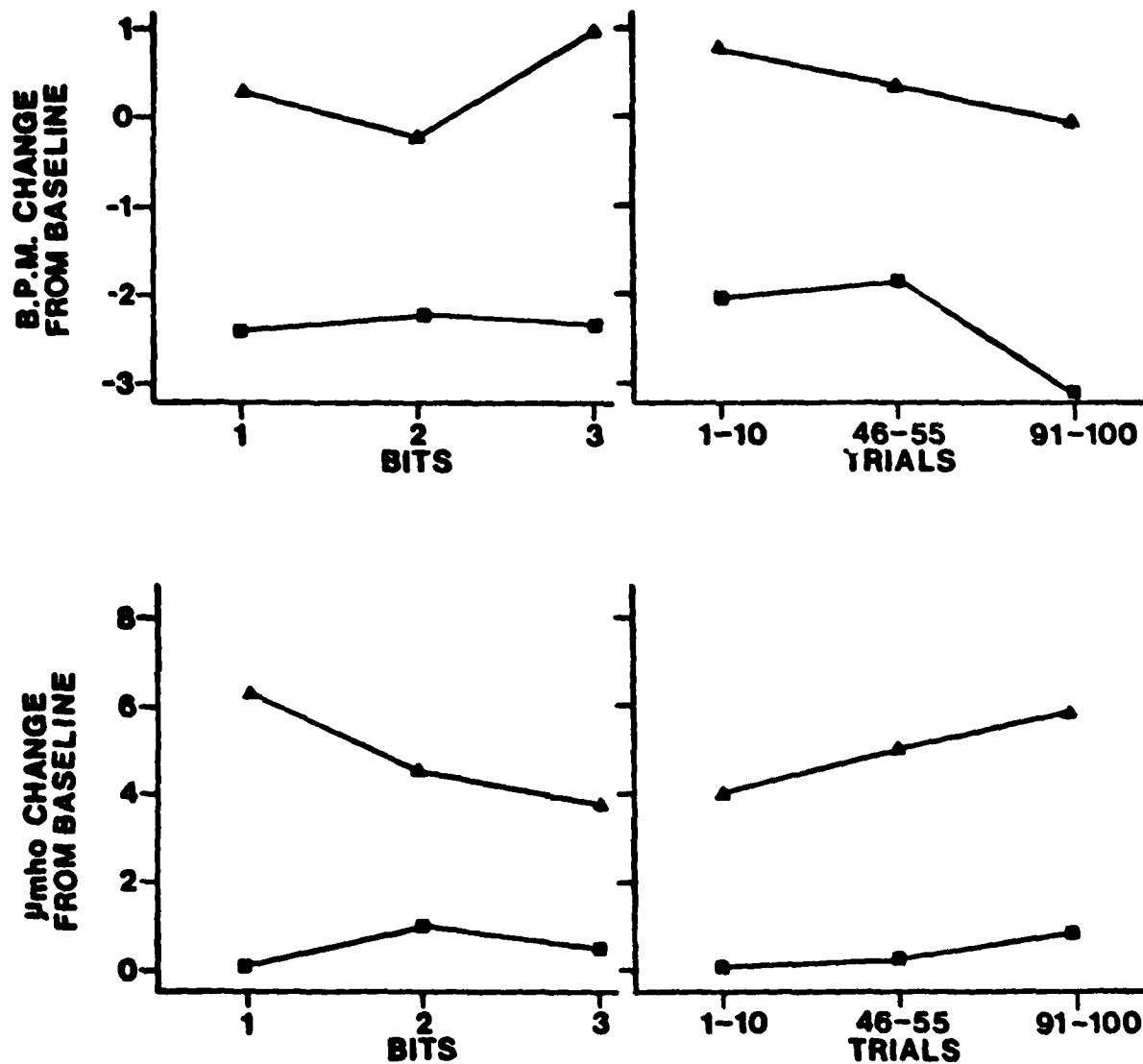


Figure 4. Heart rate (expressed in beats per minute, B.P.M.) shown in top two panels and skin conductance (expressed in μmho) shown in bottom two panels as functions of bits of information (1, 2, and 3 bits correspond to 2, 4, and 8 alternative choice, respectively) and trial blocks for the fast (triangles) and slow (squares) subjects of Experiment 2.

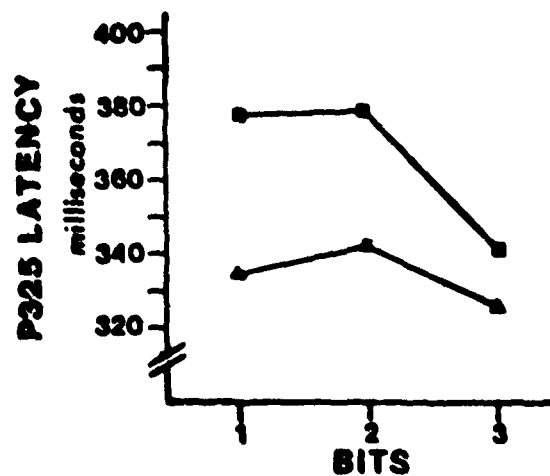
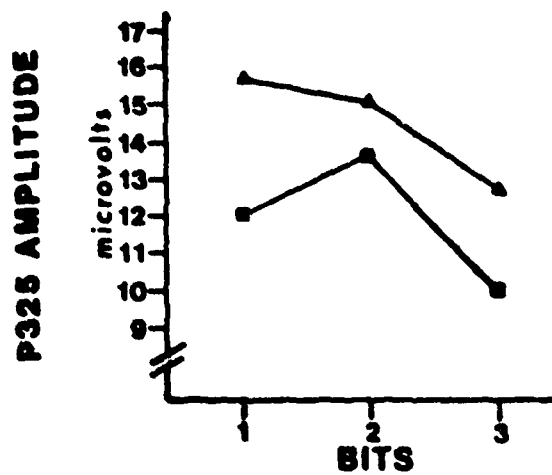
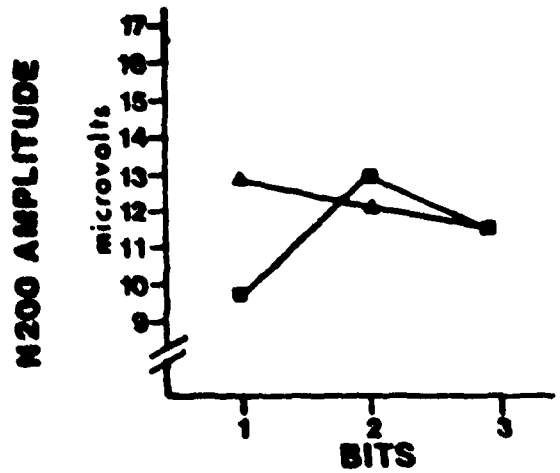


Figure 5. Vertex potentials as functions of bits of information for the fast (triangles) and slow (squares) subjects of Experiment 2.

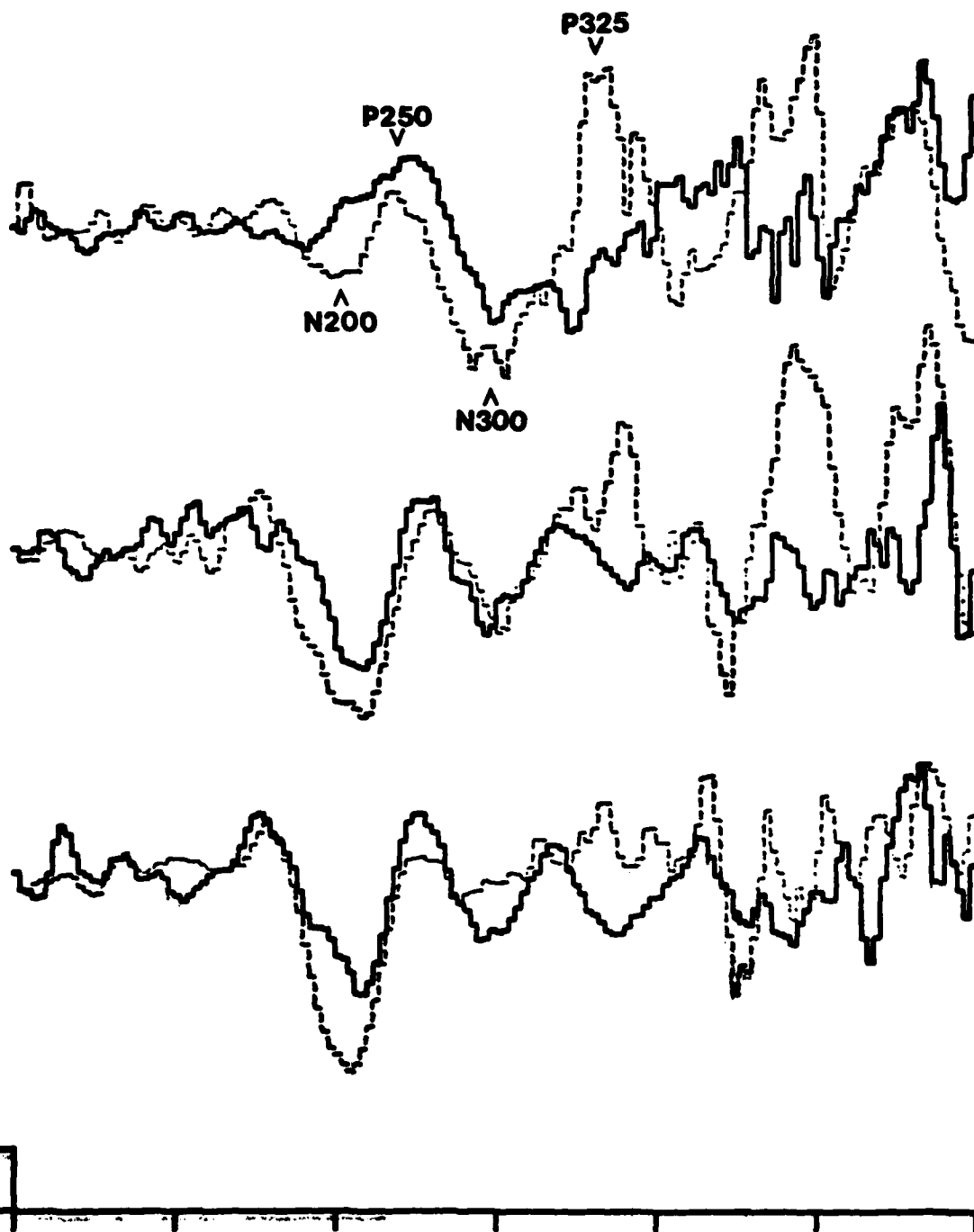


Figure 6. Left vertical mark is both calibration marker (1 μ V) and stimulus onset marker. Positive up. Abscissa ticks are 100 msec apart. Light traces are vertex potentials for fast (Q1) RTs, heavy traces, slow (Q4) RTs. Top traces, Condition 1, middle traces, Condition 2, and bottom traces, Condition 3.

that the variance of all components increased markedly in Condition 3 with the greatest increase in variance observed in the P325 component. This increased variance resulted in fewer significant differences as task difficulty increased: In Condition 1, N200 amplitude was larger and N200 latency shorter on fast trials ($t(5) = 2.89$ and 9.30). In Condition 2, only the P325 component differed between fast and slow trials. The amplitude was greater for fast trials ($t(5) = 2.57$) and the latency shorter ($t(5) = 3.16$). Finally, in Condition 3, there were no reliable differences between fast and slow trials for any of the vertex components.

Correlational analyses-Within-subject correlations revealed reliable relationships between P325 amplitude and RT across Conditions (average $r = -.43$). Also, P325 latency correlated reliably with RT in Condition 1 ($r = .29$), approached the .05 level in Condition 2 ($r = .19$), and dropped to essentially zero ($r = .07$) in Condition 3. Between-subject correlations showed the same pattern for RT and P325 amplitude (average $r = -.42$, not reliable with 4 df), but the correlation of RT and P325 latency was essentially zero (average $r = -.08$).

Summary-The pattern of results was generally the same as for Experiment 1. Between subject comparisons revealed consistent but not statistically reliable augmentation of HR, SCR, and vertex potentials for fast, compared with slow, subjects. An exception was that P325 amplitude was significantly greater for fast, compared with slow, subjects over Conditions. The correlations suggested that fast RTs were associated with high amplitude P325s, but not reliably so. The results of the within comparisons revealed reliable augmentation of vertex potentials for fast trials in Condition 1, in agreement with the results of Experiment 1. However, as information load increased, fewer differences were seen; in Condition 2, only P325 amplitude and latency were different for fast and slow trials, and in Condition 3, no reliable effects were detected. Also in agreement with the results of Experiment 1, the correlations indicated that fast RTs were associated with short latency, high amplitude P325s, but the RT-P325 latency relationship deteriorated as information load increased.

VI. Experiment 3

Introduction

Experiments 1 and 2 dealt with choice reaction time which is among the simplest behavioral tasks in which information load can be systematically varied. In the third and fourth experiments, a more complex task was investigated, which involved symbol recognition (letters of the alphabet) and

required subjects to make semantic judgments. This is the letter matching task which is described in detail by Posner and Boies, (1971). Briefly, subjects are presented with a pair of letters which may be physically identical (e.g., AA), identical in name only (e.g., Aa), or different (e.g., AB). The subject has two response keys, marked "same" and "different" and is instructed to make a judgment and respond as rapidly as possible. Posner and others have shown repeatedly that reaction time to name-identity pairs is considerably longer than to physical identity pairs.

The condition just described is the "same-different" condition. In another condition, referred to as the "vowel-consonant" (VC) condition, the task is changed in only one way: the letter pairs might be both vowels (e.g., AE), both consonants (e.g., BD), or one of each. In this case, the correct response would be "same" for the first two examples, since both are vowels or consonants, and "different" for the case when one is a vowel and one a consonant. Again, the literature is consistent in showing that RTs are longer to vowel and consonant matches than to either physical matches or name matches (Posner & Mitchell, 1967).

The rationale underlying the RT differences in the letter matching task is that physical matches require minimal central processing since decisions can be made on the basis of stimulus characteristics alone (that is, in a physical match, all the lines used to draw the two stimuli are identical). However, in the name-identity match, a semantic translation is required which requires more central processing time (that is, lower and upper case letters do not look alike, but they share the same name). Finally, the vowel-consonant matches are the most difficult since they are "same" only in the sense that they are members of the same abstract sets, and this requires even more central processing time.

The first two experiments utilized choice RT paradigms, which minimized semantic translations. The letter matching task represents a higher level of stimulus processing since correct responses depend on the subjects' ability to manipulate abstract symbols. In this manner, the letter matching task represents a level of task difficulty a step above that of the choice RT task and serves as the next logical step in investigating the relationships among psychophysiological and behavioral indices of performance.

Method

Subjects-Five of the six subjects of Experiments 1 and 2 also served in this experiment. The sixth left the university and

was replaced by another of approximately the same age and background.

Apparatus-The letters were presented on a Hewlett-Packard 2948A Graphics Terminal. The letters used in both the same-different (SD) and vowel-consonant (VC) conditions were upper and lower case, A, I, D, and F, drawn with the aid of the 2948A's optional large letter set. As in Experiments 1 and 2, the PDP 11/34 system was programmed to present all stimuli, record all data, and time the various intervals of the experiment. The "same" and "different" response keys were simple telegraph keys purchased at a local electronics supply store. Both were mounted horizontally to the immediate right of the display. The keys were 7.5 cm apart.

Procedure-Subjects served in two sessions. During the first (practice session), the task was carefully explained and each subject received 96 practice trials on the SD condition and, after a 10 min break, 96 practice trials on the VC condition. Each trial started with the appearance of a 2.5 by 2.5 cm checkerboard which appeared in the center of the display and served as a fixation point. After 4 sec, the checkerboard was replaced by two letters side by side, and these remained visible until the subject responded to either the same or different key. The letters subtended slightly less than 2 degrees of arc and were therefore foveal.

When the subject responded, the letters disappeared, and a 4 sec intertrial interval was imposed, during which time the display screen was blank. The beginning of the next trial was again marked by the appearance of the checkerboard. The length of this first session was approximately 1 hour.

The second session, which followed the first by 1 to 3 days, was the same except that biopotential leads were attached as described for Experiment 1, and 192 trials were administered for each of the SD and VC conditions. Again, conditions were separated by a 10 min break.

Within each sequence of 48 trials, there were equal numbers of "same" and "different" letter pairs. All letters appeared equally often, and the positions of upper and lower case letters were balanced.

Results and Discussion

Behavioral Results-The median RT to physically identical stimuli (PI) was 542 msec averaged across the six subjects. The comparable figure for name-identity (NI) was 602 msec. This difference was reliable ($p < .05$), which replicates

earlier findings that physical matches require less processing time than name matches. In the VC condition, the median R1 to VC matches was 709 ms, considerably longer than to P1 or N1, which is also in agreement with previous results (Posner & Mitchell, 1967; Posner & Boies, 1971).

Psychophysiological results-Although between-subject comparisons were made, neither HR, SCR, nor specific components of the vertex potential differed for fast versus slow subjects in spite of the fact that substantial R1 differences were present. Concerning within-subject comparisons, HR was stable throughout the experiment and was not different for the P1 versus N1 matches nor for the Q1 versus Q4 R1s. Skin conductance did reliably increase over trials ($p < .05$). This function (not shown) was linear and increased monotonically from the first to the last trial block, indicating a small but steady increase in sympathetic activity as the experiment progressed.

Figure 7 displays the vertex potential averaged over subjects. The three superimposed traces represent the vertex potential for P1, N1, and VC matches. Recall that the median R1s were 542, 602, and 709 msec, respectively, for these three matches and note that the peak amplitude of the prominent positive component occurring approximately 160 to 200 msec following stimulus onset (P200) increases monotonically with median R1. These amplitude differences were reliable for the P1-N1 comparison ($p < .05$) and the VC-P1 comparison ($p < .05$).

In summary, the behavioral results agreed with those previously published by others (Posner & Boies, 1971): R1s were shortest for physical matches, intermediate for name matches, and longest for vowel-consonant matches, which presumably is indicative of the greater amounts of central processing time required to decide that two stimuli have the same name or belong to the same abstract sets (vowels and consonants). HR remained stable throughout the experiment and did not vary as a function of condition, letter type, or speed of R1. SCR did increase reliably over trials suggesting a small but consistent increase in sympathetic arousal. The vertex potential results were quite interesting. The peak amplitude, averaged over subjects, of a P200 component bore an orderly relationship to task difficulty as assessed by R1: the amplitude was smallest for the easiest match (P1), intermediate for the match of intermediate difficulty (N1), and largest for the most difficult match (vowel-consonant). This result is important since the P200 occurs several hundred msec before the behavioral response and could, therefore, be a reflection of central decision-making processes.

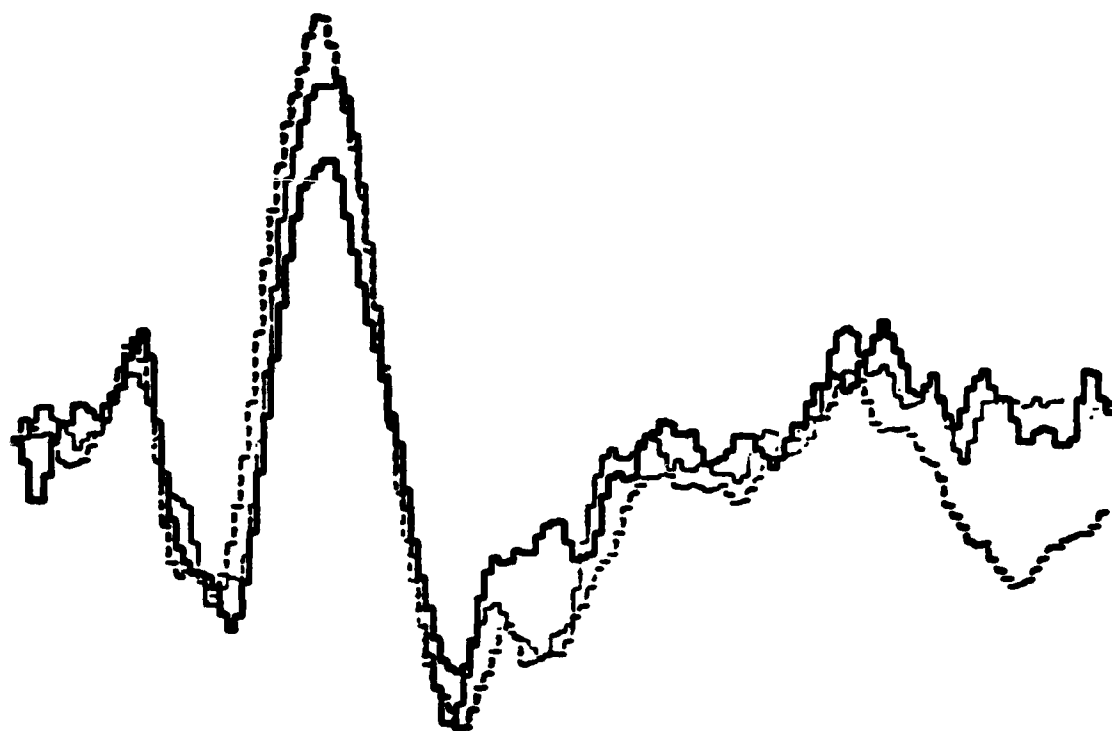


Figure 7. Left vertical mark is both calibration ($1 \mu\text{V}$) and stimulus onset marker. Positive up. Abscissa ticks are 100 msec apart. Heavy solid trace is vertex potential for PI matches, light solid trace is vertex potential for NI matches, and broken trace is vertex potential for VC matches.

Note that, in the averaged waveforms (Figure 7), there are no clearly defined late components. Inspection of individual trial data, however, showed that a positive component was frequently present in the latency range of 400 to 600 msec, but variability of latency was extreme both within and between subjects. Thus, when trials were averaged, no clearly defined peak emerged. New techniques permit correcting the latency of peaks when variance is high, but this laboratory is not currently prepared to apply these techniques. This topic is elaborated in the General Discussion.

VII. Experiment 4

Introduction

The letter matching task has been investigated in many laboratories with numerous modifications of the basic methodology. Experiment 3 employed one common version in which the two letters are presented simultaneously. In another common version, investigated here, the second letter is presented some number of msec following the first. RTs are generally faster in the latter, compared with the former, version of the task presumably because the subject has time to encode the first letter during the interstimulus interval (ISI). Thus, by the time the second letter appears, half of the information relevant to making the same-different judgment has already been encoded and RTs following the presentation of the second letter are correspondingly faster (Posner & Boies, 1971).

Using this paradigm with an ISI of 1 sec, Posner, Klein, Summers, and Buggie (1973) recorded vertex potentials and obtained waveforms which differed in many respects from the ones obtained here in Experiment 3 using simultaneous (ISI=0) presentation. Specifically, they report augmentation of late components (i.e., P300 for "different" compared with "same" matches). Since it appeared important to determine whether the extreme variability of late components in Experiment 3 was due to this single difference in methodology (specifically, changing the ISI from zero to some value larger than zero), Experiment 4 was conducted. This experiment is identical to the same-different condition of Experiment 3 except that ISI is set at 900 msec.

Method

Subjects-The same subjects that served in Experiment 3 served in this experiment.

Apparatus-The equipment used was identical in all respects to that of Experiment 3, including the specific letter set used.

Procedure-The procedure was identical to Experiment 3 except that the second letter followed the first by 900 msec. Since the appearance of the first letter acted as a warning stimulus, the checkerboard fixation point was considered superfluous and was eliminated. Following the method of Posner et al. (1973), the first letter remained visible for 250 msec, and the second letter remained visible until the subject responded.

Results and Discussion

Behavioral results-The median RT to PI was 377 msec and 438 msec to NI. This difference was reliable ($p < .05$) which is in agreement with earlier results indicating that physical matches require less central processing time than name matches. Also, RTs here are considerably shorter (about 150 msec shorter) than in the PI-NI matches of Experiment 3. This, too, is consistent with the Posner and Boies (1971) results and presumably demonstrates that subjects process and encode half of the task relevant information during the 900 msec ISI.

Psychophysiological results-As in Experiment 3, between-subject analyses failed to yield any consistent differences between fast and slow subjects. Also in agreement with the results of Experiment 3, within-subject analyses revealed that HR was stable throughout the experiment and did not vary as a function of PI versus NI, Q1 versus Q4 RTs, or trials. The SCR results were also equivocal. The data for one subject was lost due to a disk directory error. Of the remaining five subjects, four showed increases in SCR from the first to last trial block, and one subject showed a slight decrease. The means showed an increase of about 2 umho from the first to last trial block, a trivial increase.

The vertex potential results were quite interesting and are represented in Figure 8 where PI and NI are plotted separately, averaged over the six subjects. There are several points of interest. First, the vertex responses to the first letter of the pair (top trace) are essentially identical for PI and NI. This is to be expected since the subject is simply encoding the letter and does not yet know whether a PI or NI match is to be made. However, the vertex potential shows clear and reliable PI-NI differences as early as 150-200 msec following appearance of the second letter of the pair (lower trace). Three major components were identified in this analysis: a positive component occurring about 200 msec following stimulus onset (P200), a negative at 260 msec (N260), and a second positive at 370 msec (P370). The latencies of all three components were shorter ($p < .05$) for the easy (PI) match compared with the harder (NI) match, and the peak-trough

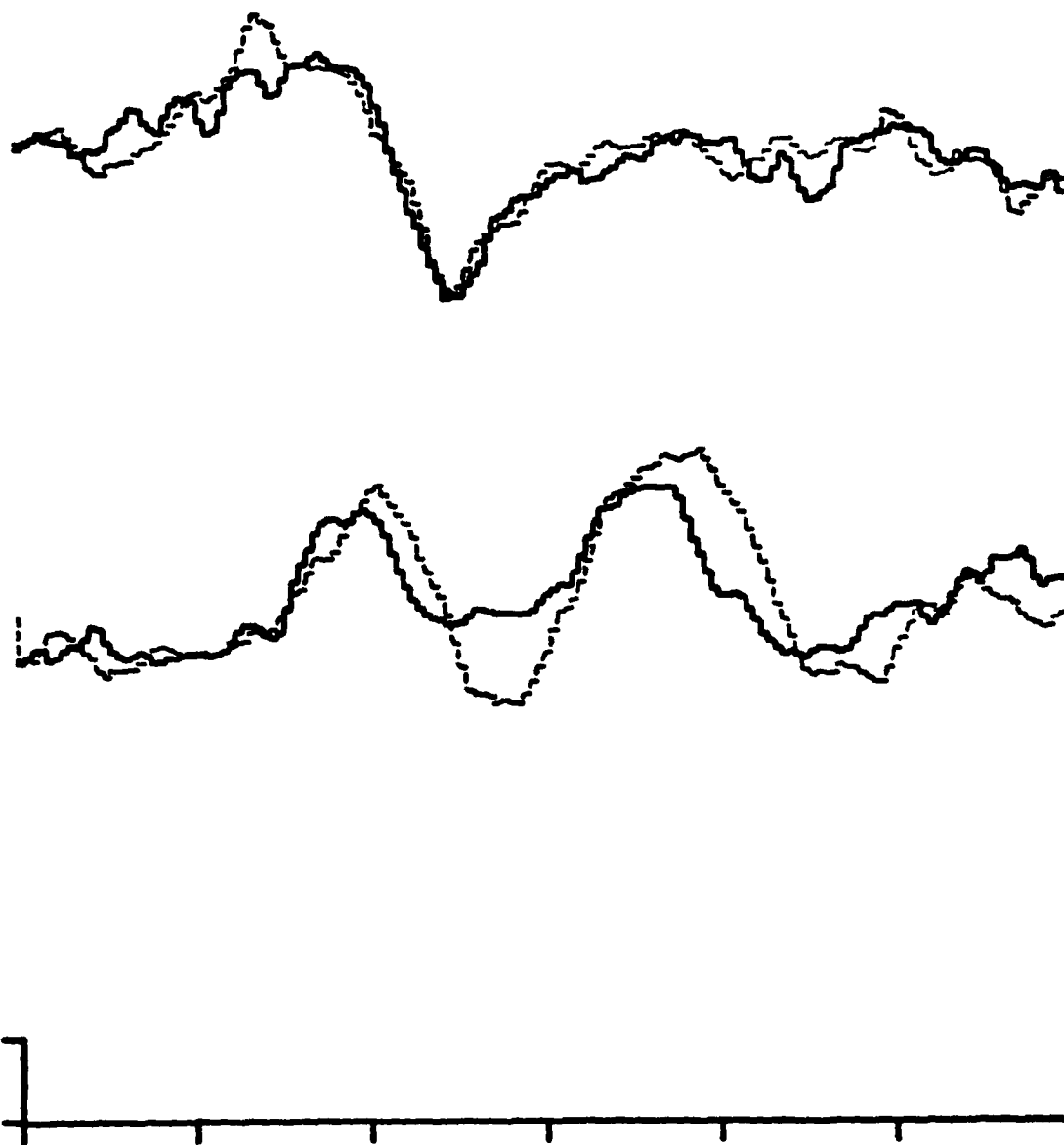


Figure 8. Left vertical mark is both calibration ($1 \mu V$) and stimulus onset marker. Positive up. Abscissa ticks are 100 msec apart. Top traces, vertex potential responses to first letter of pair. Bottom traces, vertex potential responses to second letter of pair presented 900 msec after first letter. Light traces are NI matches; heavy traces, PI matches.

amplitude differences (P200-N260 and N260-P370) were reliably ($p < .05$) greater for NI than for PI. In simple terms, vertex potentials were larger in amplitude and occurred later in time for NI compared with PI suggesting that increases in latency and amplitude of the vertex components are indicative of greater task difficulty.

VIII. General Discussion

The results of the four experiments showed that certain components of the vertex potential were consistently related to performance and task difficulty, while the indicators of sympathetic arousal (HR and SCR) tended to be task specific: Changes in HR and SCR were more apparent in the choice RT tasks than in the letter matching tasks. This appeared to be due to the different motor responses required and the amount of subject involvement in the two types of tasks. The choice RT paradigm required movement of the entire right arm while the letter matching task required movement of only one finger. Subjects volunteered their opinions that both tasks were boring and that the letter matching task was particularly tedious. The choice RT task was less tedious because of the physical involvement of hitting the response button as rapidly as possible (and rather violently for some subjects) which took on some aspects of a game. The reduced HR and SCR responsivity in Experiment 2, compared with Experiment 1, was probably due to the fact that the same subjects served in both experiments and they were habituating to the task, finding the "game" less interesting with repeated exposure. These observations can be used to make inferences concerning the extent to which subjects are involved in a task. In Experiments 3 and 4, HR remained stable and SCR increased a small amount. This pattern of responses is indicative of low sympathetic arousal, and subjects reported that these two tasks were tedious. Subjects reported that Experiment 1 was less tedious and it was in this Experiment that clear evidence was found for increased sympathetic arousal for fast, compared with slow subjects. In a gross sense, then, task involvement was related to sympathetic arousal, and refinements of this approach could be used to scale task difficulty in a manner which takes into account the subject's perception of whether the task was tedious (low arousal) or interesting and involving (high arousal).

The waveform of the vertex potential was different for different stimuli. Simple light onset against a dark background (choice RT tasks) evoked an early negative component (N200) followed by late components, while complex shapes (upper and lower case letters in the letter-matching tasks) evoked an early positive component (P200) followed by late components.

Such differences have been noted in prior literature and appear to be related to stimulus complexity. For behavioral research, the polarity of the component is not as important as whether the latency and/or amplitude varies systematically with psychological dimensions.

Task difficulty varied less in the choice RT tasks than in the letter-matching tasks, therefore, analyses of the former data focused on small differences in RTs within and between subjects. In general, the main components of the vertex potential were shorter latency and/or higher amplitude when RTs were short, but these relationships became less clear as task difficulty increased since the variance of EP component increased. When analyses focused on larger differences in task difficulty (the letter-matching tasks), an early, P200 component, increased in amplitude with increasing task difficulty. Less clearly, and only with the 900 msec ISI, the late component (P300) also showed higher amplitude and longer latency for the more difficult tasks. In general, then, the vertex potential discriminated task difficulty and performance as well or better than in previously published research using simpler behavioral paradigms.

In tasks where RT was longest (Condition 3 of Experiment 2 and the zero ISI condition of the letter-matching task), variance of the vertex potential increased to the point where meaningful analyses were precluded. Other authors have encountered this problem and have derived various solutions. One of questionable worth is to have judges examine the record of each trial and discard trials on which a "clear" P300 cannot be defined. This method can result in discarding perhaps 50% of the data collected but does result in substantial correlations between late components and RTs (Ritter, Simpson & Vaughn, 1972). Another approach is to use a wave subtraction procedure in which the EP to a "no signal" trial is subtracted from the EP to a "signal" trial. This tends to enhance any EP component which is uniquely related to signals (Goodin et al., 1978) and appears to be a justifiable procedure. The limitation is that the experiment must be specially designed to present signal and no signal trials and such contrived designs might have little generality to real world situations. A third solution is to construct an empirically derived template of prototypical P300 and use reiterative computer programs to find the best fit between the template and the P300 for any given trial. Kutas, McCarthy, and Donchin (1977) have described one possible method for accomplishing this. It has the advantage of not discarding trials but rather computing an index of how well a component on any given trial conforms to the prototypical P300s. Another advantage is that contrived designs, necessary for the waveform subtraction procedure, are not necessary for

prototypical matches. Clearly, the prototype matching concept can be applied to components other than the P300 so that all major components of the vertex potential could be analyzed in this manner.

This laboratory is currently developing software to accomplish template matching, but this effort is not yet completed. Therefore, the approach used was to define each subject's score as the maximum amplitude deflection occurring within the latency window typified by the averaged waveform. No trials were discarded. While this decision rule was adequate to detect differences in several important comparisons, it almost certainly was not an optimal decision rule since it assumes that the "important" component is the one of highest amplitude. To the extent that this assumption is erroneous, the decision rule would wrongly identify components. This in turn would lead to low magnitude correlations and failure to detect mean differences when attempts were made to relate a specific EP component to performance. It seems likely, therefore, that the differences reported here could have been enhanced by application of a more sophisticated decision rule.

Summary and Conclusions

This was a preliminary effort to screen the usefulness of several psychophysiological measures as indicators of human performance and task difficulty. Few subjects were used; the behavioral tasks were simple, but not as simple as the paradigms employed in the majority of earlier psychophysiological experiments. Given the preliminary nature of the effort, the outcomes were quite encouraging. Heart rate and skin conductance were used as indicators of autonomic arousal or activation. Both increased when the task was involving (by subject reports) and failed to increase when the task was not involving. Selected components of the vertex potential were augmented (short latency and/or high amplitude) when subjects performed well and/or task difficulty was high. Variance problems were encountered in some instances, but this is not considered as a criticism of the psychophysiological approach; rather, it underlines the need to develop sophisticated analyses to decode sophisticated data.

References

Duncan-Johnson, C. C. & Donchin, E. On quantifying surprise: The variations of event-related potentials with subjective probability. Psychophysiology, 1977, 14, 456-467.

Edelberg, R. Electrical activity of the skin: Its measurement and uses in psychophysiology. In N.S. Greenfield and R.A. Sternbach (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart, & Winston, 1972.

Goodin, D.S., Squires, K.C., Henderson, B.H., & Starr, A. An early event related cortical potential. *Psychophysiology*, 1978, 15, 360-365.

Gunn, C.G., Wolf, S., Block, R.I., & Person, R.J. Psychophysiology of the cardiovascular system. In N.S. Greenfield and R.A. Sternbach (Eds.), *Handbook of Psychophysiology*. New York: Holt, Rinehart, & Winston, 1972.

Haider, J., Spong, P., & Lindsley, D.B. Attention, vigilance, and cortical evoked potentials in humans. *Science*, 1964, 145, 180-181.

Karlin, I. Cognition, preparation, and sensory-evoked potentials. *Psychological Bulletin*, 1970, 73, 122-136.

Kutas, M., McCarthy, G., & Donchin, I. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, 1977, 197, 792-795.

Lacey, B.C. & Lacey, J.I. Two way communication between the heart and the brain. *American Psychologist*, 1978, 33, 99-133.

Posner, M.I. & Boies, S.J. Components of attention. *Psychological Review*, 1971, 78, 391-408.

Posner, M.I., Klein, R., Summers, J., & Buggie, S. On the selection of signals. *Memory and Cognition*, 1973, 1, 2-12.

Posner, M.I. & Mitchell, R.F. Chronometric analysis of classification. *Psychological Review*, 1967, 74, 392-409.

Ritter, W., Simpson, R., & Vaughan, H.G. Association cortex potentials and reaction time in auditory discrimination. *Electroencephalography and Clinical Neurophysiology*, 1972, 33, 547-555.

Ritter, W., Simpson, R., Vaughan, H.G., & Friedman, D. A brain event related to the making of a sensory discrimination. *Science*, 1979, 203, 1358-1361.

Roman, J., Older, H., & Jones, W.I. Flight research program: VII. Medical monitoring of Navy career pilots in combat. *Aerospace Medicine*, 1967, 38, 133-139.

Spong, P., Haider, M., & Lindsley, D.B. Selective attentiveness and cortical evoked responses to visual and auditory stimuli. *Science*, 1965, 148, 395-397.

Sutton, S. The specification of psychological variables in an averaged evoked potential experiment. In E. Donchin and D.B. Lindsley (Eds.), Averaged Evoked Potentials: Methods, Results, and Evaluations. NASA SP-191, Washington, DC: U.S. Government Printing Office, 1969, Pp 237-297.